

DOCUMENT RESUME

ED 178 552

TM 008 993

AUTHOR Chapman, Paul Davis
TITLE Schools as Sorters: Testing and Tracking in California, 1910-1925.
PUB DATE 8 Apr 79
NOTE 28p.; Paper presented at the annual meeting of the American Educational Research Association (63rd, San Francisco, CA, April 8-12, 1979)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Ability Grouping; Adaption (Ideas); Educational Change; Educational Philosophy; Educational Research; Educational Researchers; Elementary Secondary Education; Flexible Progression; *Group Intelligence Testing; Immigrants; Intelligence Differences; *Intelligence Tests; Public School Systems; Social Values; *Student Testing
IDENTIFIERS *California; California (Oakland); California (Palo Alto); California (San Jose); Dickson (Virgil); Proctor (William V); *Terman (Lewis M); Young (Kimball)

ABSTRACT

This paper explores the reasons for the rapid adoption of intelligence tests by the public schools, and the historical relationship between testing and ability grouping or tracking. Case studies are presented of three California communities--Oakland, San Jose, and Palo Alto--between 1910 and 1925. These communities have been selected because they were pioneers in the use of intelligence tests; because they were associated with Lewis Terman and Stanford University; and because they were representative of urban communities in terms of size, socioeconomic status, and ethnic diversity (including large immigrant populations). Reasons for the rapid emergence of intelligence testing are suggested, followed by a review of Lewis Terman's work and philosophy. The case study of the Oakland schools describes the efforts of John W. McClymonds, Vinnie Hicks, Ellwood P. Cubberley, Fred Hunter, and Virgil Dickson. The San Jose case study describes the work of Kimball Young and Walter L. Bachrodt. The testing done by William V. Proctor is described in the case study of Palo Alto. These histories are said to support the suggested reasons for the rapid adoption of intelligence testing: testing technology, educational problems of the time, and social values which supported efficiency in tracking. (GDC)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

"Schools as Sorters: Testing and Tracking in California, 1910-1925"

Paul Davis Chapman
School of Education
Stanford University

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Paul Davis Chapman

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Paper Presented at the Annual Meeting of the American
Educational Research Association, San Francisco

April 8, 1979

Session Number 4.22

Copyright © Paul Davis Chapman 1979

In the first quarter of the twentieth century American schools witnessed a dramatic change in the means by which students were classified for instruction. In the first decade of the century, psychologists on both sides of the Atlantic developed experimental tests for the purpose of classifying students by mental ability. Over the next fifteen years these tests were adopted with astonishing swiftness for use in the public schools. By 1925 psychologists had developed over seventy-five tests of general mental ability; according to some estimates yearly test sales approached four million by mid-decade. During these same years the practice of ability grouping became widespread. A survey conducted by the Bureau of Education in 1926, for example, revealed that a large majority of urban schools classified students according to ability. This same study also showed that 86 percent of the cities in the sample used intelligence tests as one basis of classification.¹

The purpose of this essay is to explore the origins of the use of intelligence tests for the purpose of classifying students into ability groups. The paper seeks answers to two related questions. First, why was this educational innovation--the intelligence test--adopted so rapidly in the public schools? Second, what was the historical relation between testing and ability grouping, or tracking? Part of a larger study of Lewis Terman and the intelligence testing movement, this essay will examine the adoption of tests in three California cities: Oakland, San Jose, and Palo Alto. These case studies were chosen for several reasons. Each pioneered in the use of tests around the time of the First World War. Each had a direct link to Stanford University, where Lewis Terman was developing a series of intelligence and achievement tests that were to become the most widely used in the nation. Together these cities represented a distribution of urban communities in terms of size, ethnic diversity, and socio-economic make-up. This study of the testing movement explores the relationship between universities and school systems. On the one hand, it examines the role played

by professional psychologists in developing intelligence tests and prescriptions for their use; on the other hand, it explores both the response of the schools to internal needs as well as the psychologists' plans for reform.

To explain the rapid emergence of intelligence testing and ability grouping, three principal arguments will be suggested. First, a network of newly-emergent professionals--university professors of psychology and public school administrators--advocated the widespread use of tests to classify students. They were supported in their efforts by philanthropic foundations, educational publishers, and such national institutions as the Bureau of Education. Second, intelligence tests were adopted in the schools as a result of the changing structure and function of the schools. At the turn of the century, schoolmen faced a vast array of problems. Urban school enrollments mushroomed as a consequence of massive immigration, the move from the country to the city, and newly enforced compulsory education laws. School populations represented, as well, an increasingly diverse patchwork of different ethnic groups. At roughly the same time, schools evolved a new set of objectives; the common school curriculum of the nineteenth century gave way to "life adjustment."² In response to these changes the school curriculum became much more differentiated. School administrators needed a means of classifying students in order to assign them to different curricular tracks. The intelligence tests provided such a means. Third, intelligence tests were adopted in the public schools because the use of tests reflected many of the values of the Progressive Era, the most important of which were: efficiency and conservation; the belief in science and the authority of the professional expert; the concern with individual differences; and the nativistic fear of the new immigrants. To a large degree the experience of California in the first quarter of the twentieth century supports this interpretation of how and why intelligence tests were adopted in the schools.

Terman's Prescriptions for Reform

The spearhead behind the use of intelligence tests in California, indeed a principal leader in the national movement for intelligence testing, was Lewis Terman. His development of individual and group intelligence tests at Stanford, his prescriptions for their use in schools, and his role in national organizations concerned with testing all served to make Terman a key figure in the adoption of intelligence tests by California schools. Terman's strong positions on such issues as the constancy of the I.Q., the relative role of heredity in intelligence, and the interpretation of inter-group differences in performance of the intelligence tests served as well to make Terman a key figure in the controversy over testing in the 1920's. Terman's interest in the concept of intelligence and his plans for the use of tests can be traced to his early years in education.

Born on a farm in southern Indiana in 1877, Terman entered teaching at the young age of fifteen, became the principal of the local high school, and then in 1903 travelled east to Clark University to begin work on his Ph.D. under G. Stanley Hall. Introduced to the intelligence studies of such Europeans as Wilhelm Wundt, Francis Galton, and Alfred Binet and to the work of James McKeen Cattell at Columbia, Terman investigated, as he later recalled, the psychology of genius, the measurement of intelligence, and the phenomena of individual differences. In his dissertation on Genius and Stupidity, he examined "deep and fundamental lines of cleavage among individual mentalities...."³

Four years after receiving his doctorate, Terman was called to Stanford University in 1910 to help develop the Department of Education, and it was here that he quickly established himself as a national figure in the field of intelligence testing. Upon arriving at Stanford, Terman began a revision of Binet's intelligence scale, a project that would culminate six years later in the publication of the Stanford-Binet Intelligence Test. Terman translated

Binet's tests into English, extended the scale at its lower and upper levels (approximately five and fourteen), and expanded the number of tests given at each age. In addition, he standardized the results on a sample of individuals he believed to be typical of the American population. Perhaps most importantly, he gave full expression to the concept of the intelligence quotient, or I.Q., though the idea had originated with Wilhelm Stern in Germany several years earlier. Terman popularized the concept of the quotient, a single number which would relate mental and chronological age and allow comparisons between individual and normal mental development.

While preparing the Stanford-Binet revision, Terman fashioned a preliminary plan for how the tests might be used in schools. Like Binet, Terman first became interested in the use of tests to identify "abnormal" children, those who were either "feeble-minded" or gifted. Reporting on his work in 1911 he predicted that "before many years it will probably become a matter of course to apply serial mental tests in the public schools to all pupils who are retarded or about to become retarded, or who give indication of unusual ability." Terman went on to predict a much wider application of tests. Familiar with Frederick W. Taylor's concept of "scientific management" and with Leonard Ayres' recent study of school retardation, Laggards in Our Schools, Terman observed that intelligence tests could be used by school administrators to modify the entire school curriculum to suit individual abilities:

Now that the individual treatment of pupils in the schools has begun, there is no stopping short of this ideal. Tests must be developed which will enable us to differentiate all degrees of intellectual ability and all kinds of intellective unevenness.

The next year Terman had his first chance to put these ideas into practice.⁴

Using the Stanford-Binet revision to survey the "mentally defective children" of the San Luis Obispo schools in 1912, Terman discovered that all twenty-four had low I.Q.'s. In his analysis of these educational misfits, Terman

did not use kind words:

"They clog the educational machinery. They consume a disproportionately large part of the regular teacher's energy. They pull down the standard of achievement for other children. Finding themselves chronic failures they either become disheartened and dejected, or else they grow case-hardened and apathetic."

Terman recommended that the superintendent, Charles Small, use the results of the survey to provide special classes for the backward children and suggested that the research provided a "good argument for the introduction of manual training, domestic science" into the schools.⁵

Over the next several years, Terman shaped these preliminary plans into a systematic proposal for the use of intelligence tests in schools. He believed that the tests would: help schoolmen to "identify the feeble-minded" and to "choose the methods and matter of education which will guarantee for such children the best possible returns for their efforts"; help teachers and administrators address individual differences by promoting students "largely on the basis of intellectual ability"; assist in the development of vocational guidance through which "it would then be possible to lay before each child at the end of this school course an array of occupations in which (as far as intelligence is concerned) he might reasonably be expected to succeed"; and finally help promote "the segregation of the feeble-minded" in order to "remove a demoralizing and retrograding influence from the lives of many normal children." With the publication of The Measurement of Intelligence, a guide to the Stanford-Binet, these proposals reached a national audience.⁶

In these same years Terman sought to win approval for his plans by demonstrating the usefulness of intelligence tests for progressive social reform. Enlisted by California Governor Hiram Johnson, who was on the 1912 Progressive Party ticket with Theodore Roosevelt, Terman investigated the efficiency of the state's welfare institutions with mental surveys of The Whittier State School for Boys, San Quentin Prison, an orphanage, as well as the schools of

---6---

an entire county. Noting that "the children of subnormal mental endowment undoubtedly constitute one of our gravest social problems," he became an early advocate of eugenics to solve the problems of crime, poverty and prostitution.⁷

When the Measurement of Intelligence was published, Terman could little realize that soon he would have a dramatic opportunity to demonstrate his ideas for the use of intelligence tests. Within a year, however, war was declared on Germany, and in May 1917 Terman joined a committee of the American Psychological Association to consider ways psychologists might aid the war effort. Basing its work largely on the group intelligence test recently created by Arthur Otis, A Terman graduate student, the committee developed the Army Alpha and over the next eighteen months tested and classified some 1.7 million recruits. The significance of the army experiment was twofold. The project brought national attention to the intelligence tests, demonstrating to the general public as well as schoolmen their practical value for classifying large numbers of people. The experiment also led to the first group intelligence tests prepared specifically for widespread use in the public schools.

In January 1919 Terman and his army colleague, Robert Yerkes of Harvard, approached the General Education Board of the Rockefeller Foundation for a \$25,000 grant for the purpose of "developing and standardizing an intelligence scale for the group examination of school children, -- a scale for the measurement of native ability." The support was quickly forthcoming, and a group of five psychologists met in Washington in March to prepare the test. Within a year the committee had completed its work; in the fall of 1920, the World Book Company printed and distributed 400,000 National Intelligence Tests for use in public elementary schools.

Terman's role in the Army experiment and in developing the National Intelligence Test catapulted him into a position of national leadership in the

quickly burgeoning movement to use intelligence tests in schools. He expanded the technology of testing, producing the Terman Group Test for grades seven through twelve, which by 1923 sold over half a million copies,⁹ and the Stanford Achievement Test, for use in his longitudinal study of gifted children. Self-appointed spokesman for the movement, he waged a rancorous debate within the psychological profession and with the media regarding the value of the tests. As President of the American Psychological Association in 1923, he argued vigorously that the intelligence test had value not only in theoretical studies of intelligence, but as a practical tool for school administration. It was in the use of intelligence tests for school reform that Terman's ideas had their most important impact.

In two major post-war studies, The Intelligence of School Children and Intelligence Tests and School Reorganization, Terman outlined the ways in which tests could be used in schools. With the slogan "a mental test for every child," he proposed a systematic program of intelligence testing. All children would be given an individual intelligence test upon entrance; from the third grade on children would be given group tests every other year. In each school system, the program could be conducted by school psychologists, directors of research, or school administrators. After testing, students would be sorted into "homogeneous class groups." Terman advocated a "multiple-track plan" that would provide for five ability groups in a typical elementary school system of 2000 as outlined below:

TABLE I. MULTIPLE-TRACK PLAN

<u>Group</u>	<u>No. in Each Group</u>	<u>Remarks</u>
Gifted	2.5%, or 50 pupils	Two classes, 4 grades in each
Bright	15%, or 300 pupils	One class per grade
Average	65%, or 1300 pupils	Four to five classes per grade
Slow	15%, or 300 pupils	One class per grade
Special	2.5%, or 50 pupils	Three ungraded classes

After classification into ability groups, students would receive different courses of study. The "slow" group, for example, would be prepared for vocational careers, while the "bright" and "gifted" groups would be prepared for advanced work in college. A system of educational and vocational guidance based on intelligence tests would steer students towards appropriate careers. "An I.Q. below 70," Terman observed, "rarely permits anything better than unskilled labor," while scores above 115 "permit one to enter the professions or the larger fields of business."¹⁰

The values implicit in Terman's prescriptions help explain why the tests were adopted so rapidly in the schools, and why, too, the testing programs evoked a sharp and outraged response from some groups. Terman's program rested upon a number of assumptions: that intelligence was largely hereditary, that the I.Q. was a relatively constant and valid measure of intelligence, and that intelligence test scores could predict with scientific accuracy future individual development. When he discovered considerable inter-group variance on intelligence test scores, and that recent immigrants and blacks scored lower than whites, Terman concluded that poor performance was largely due to inferior mental endowment. In reaching that conclusion, Terman provided scientific confirmation for the nativism and prejudice of many Americans. The values implicit in Terman's plans for school reform--especially those of science, professional expertise, and nativism--help account for the early widespread acceptance of testing, as well as the critique of tests that emerged in the middle of the 1920's. The experience of early testing programs in California cities shows how and why Terman's plans met with initial and sometimes enthusiastic support.

The Case of Oakland

In the first several decades of the twentieth century, Oakland grappled with a variety of problems--mushrooming size, increasing diversity, and soaring costs--which faced school systems across the country. The scope of change in Oakland is depicted in the following chart:

TABLE II. POPULATION GROWTH, SCHOOL ENROLLMENT AND ATTENDANCE,
OAKLAND, CALIFORNIA, 1890-1918

Year	Population by U.S. Census	Percent Increase	Total School Enrollment	Percent Increase	Average Daily Attendance	Percent Increase
1890	48,682		9,565		6,372	
1900	66,960	38	11,976	25	8,512	34
1910	150,174	124	17,621	47	12,919	52
1917	206,402	37	40,946	132	25,797	100
1918	246,519	19	48,214	18	27,304	6

The changes in Oakland's school population came with the flood of immigrants to America around the turn of the century, with the stream of refugees who fled the San Francisco earthquake and fire in 1906, and as a result of a new compulsory school law passed by the state in 1903. As Oakland's population bulged after the turn of the century, so too did it become increasingly diverse; by 1918, sixty percent of Oakland's residents were first or second generation immigrants. The rapid growth in size and diversity undoubtedly provides much of the explanation for why school expenditures increased from \$1.2 million in 1910-11 to \$2.1 million in 1916-17.¹¹

To cope with these changes, the Oakland schools introduced a series of reforms beginning around the turn of the century. In 1901 Superintendent John W. McClymonds developed a system of flexible promotions and ungraded classes. Under this system, each teacher in the elementary grades, with the assistance of the principal, divided the class "according to the ability of the pupils, as nearly as possible into two equal divisions," and promotions were then made "whenever any section or division... completed the work of the grade." In 1911 McClymonds hired Vinnie Hicks to establish a psychological clinic and to begin examining the "abnormal" children with intelligence tests. In Hicks' view, isolating the "subnormal" children--"the absent, the tardy, the sickly, the unruly, the liars, thieves and cowards"--would free the teachers to work with the normal children. Several years later, Oakland established several "special classes" to deal with the children identified by the psychological clinic. By 1915 the structure of Oakland's public school

system had become highly differentiated to meet the needs of its new constituents. The system consisted of kindergartens, intermediate schools, and several high schools including a vocational school, a technical school, two comprehensive schools, and a "university" school exclusively for college preparatory work.¹²

Like many urban school systems in the Progressive Era, Oakland came under the scrutiny of reformers who hoped to centralize the power structure and streamline the schools for efficiency. In 1914 Jesse Sears of Stanford was asked to conduct an exhaustive study of spelling efficiency in Oakland, which disclosed wide variations in achievement among the schools. The Board of Education decided to employ "a disinterested expert to investigate the reasons for the increased expense and to determine if it were possible to conduct the schools more economically without losses of efficiency." They hired Ellwood P. Cubberley, head of Stanford's Department of Education. After an intensive twelve-day investigation, Cubberley enumerated fourteen reasons for increased cost. His solutions reflected the best educational wisdom of the time: reform the school administration along corporate lines, provide the superintendent with more powerful central authority, and expand the work of the research department to provide a more scientific assessment of student abilities and accomplishments.¹³

Oakland took Cubberley's recommendations seriously and soon hired a new, aggressive superintendent, Fred Hunter. Early in his administration Hunter identified some of the "problems" facing the city's schools:

Only One Third of the boys and girls who enter Oakland High Schools complete the high school course to graduation....Almost Eight Hundred Oakland boys and girls leave school each year between Grades Six and Eight....Thirteen Thousand Oakland school children (fifty-one percent of the total number in day high and elementary schools) are overage....Sixty Per Cent of Oakland's population is Foreign Born or children of the first generation of foreign parentage....Sixteen schools in Oakland have a percentage of foreign children ranging from Fifty to Seventy Per Cent....

With this list of particulars as ammunition, Hunter launched a campaign to reform the schools to provide for individual differences, a campaign which soon

brought to Oakland's newly-founded Bureau of Research (1914) a man conversant with the latest developments in intelligence testing.

In the summer of 1917 Virgil Dickson came to Oakland with a master's degree from Stanford and extensive training in psychological testing with Terman. In his graduate study he had explored "the relation of mental testing to school administration" by using intelligence tests to place nearly 150 first-grade children entering schools in four towns near Stanford. Drawing from his Stanford experience, he came to Oakland with a three-pronged plan for testing the work of the schools: retardation studies, achievement testing, and psychological examinations.¹⁵

In September he began a systematic study of school progress and retardation in all forty-five elementary schools. The results showed less than half the students were making normal progress; in other words, a substantial percentage of Oakland students had failed one or more terms. Furthermore, the study revealed that the highest rates of failure were concentrated in the immigrant schools, as the following chart shows:

TABLE III. SCHOOL PROGRESS IN
OAKLAND ELEMENTARY SCHOOLS, 1917

Type of Schools	Average Rate of Progress		
	Rapid	Normal	Slow
Five Schools with Greatest Percentage of Foreign Parentage (60-100%)	2%	41%	53%
Five Schools with Smallest Percentage of Foreign Parentage (20-30%)	26%	42%	30%

Over half the students in schools with the highest percentage of foreign parents had failed one or more grades, while under a third of the students in the schools with the smallest percentage of foreign parents had failed. Significantly, almost a quarter of the students in the latter schools had skipped one or more terms; virtually none of the students (two percent) had skipped in the immigrant schools. For Dickson, the indictment against the schools was clear. "These figures furnish evidence to convict the schools of two things: (1) the ability to discover weakness.

in the child and hold him back. (2) the inability to discover strength in the child and move him ahead."¹⁶

Dickson then began to analyze the abysmal record of the schools with the help of intelligence tests. In December he administered to about 2000 students the group tests Terman was preparing for use in the army. In May 1918, he secured permission from his fellow graduate student, Arthur Otis, to use his recently refined group test with some 1600 students in the eighth and ninth grades. Throughout the year he tested over 3000 pupils with the Stanford-Binet, concentrating on first graders.

The intelligence testing confirmed Dickson's belief that many students were failing because they "had inherent mental tendencies that make the ordinary course of study either impossible or impractical of attainment." His plan called for "segregation" of "mentally superior" for accelerated work, and the "mentally deficient" for special class work. For those with ability who have "fallen behind" and those who "work very slowly," he recommended "opportunity classes." For those "not yet mentally old enough" for first grade, he called for "special receiving classes." Additionally he foresaw an obvious place for intelligence tests in developing a more "scientific" vocational guidance. Like Terman, Dickson also saw important social consequences of the reform of school curricula through intelligence testing. When students were faced with expectations that exceeded their abilities, he noted, the "natural result" was "loss of interest, a loss of self-respect or a resort to subterfuge...." "Social unrest, sham, and the I.W.W. spirit," he observed, "may easily have their beginnings in these early social problems...."¹⁷

Dickson's program for testing and classification met with hearty approval among administrators and teachers in the Oakland schools. A district administrator called the experiment "the most important factor in effective educational administration that has been introduced in years," while another called for tests to be given to all primary children so that they might be classified according to

ability. The principal of the Oakland High School, C.E. Keyes, reported that intelligence tests had been used to create ability groups and to make recommendations for college as well. To Dickson, Keyes expressed the hope that "before very long the classification of the school will be based almost entirely on the results obtained from the testing." W. M. Greenwell and G.E. Mortensen, principals of two elementary schools which were the largest and which also had the highest concentration of immigrants, both strongly supported the testing program. They reported that the ability groups based on tests reduced discipline problems and dropouts and that teachers and pupils alike seemed "happier." Teachers reported a "deeper and more sympathetic understanding of the child" and a new respect for individual differences. One said the "tests have thrown floods of light on problems that have heretofore baffled me," while another reported that they had given her "greater confidence in my judgment" and the "courage of my convictions to take action." Though these testimonies were solicited by Dickson for his Stanford doctoral dissertation, they do provide fragmentary evidence that school administrators and teachers both found intelligence tests of immense practical value.¹⁸

With the post-war return to normalcy, Dickson proceeded with his experimental use of tests. By the end of the 1919-20 school year, he reported that 20,000 students had taken intelligence tests, nearly half the students enrolled. That year a committee composed of Dickson, Hunter, the supervisor of the primary grades, and three elementary teachers was established to monitor the initial experiment in ability grouping based on tests. The committee recommended that the experiment be extended system wide, and Dickson noted in June 1920 that the program was "being gradually introduced into the schools of Oakland wherever the conditions permit...." At the same time, Dickson issued a cautionary warning that the group mental test was not an "infallible guide" but merely an "important tool" which could be "used to advantage in the study and placement of groups and individuals."¹⁹

Despite Dickson's warning, the results of the early experiments suggest that test results discriminated against certain immigrant groups. Dickson reported an early experiment using intelligence tests in two elementary schools where two-thirds of the children were first or second generation immigrants from Portugal, Italy, Austria, Assyria, Spain, Greece, England, and the Balkans. Recent immigrants, especially those from Southern and Eastern Europe, generally tested below average on the intelligence tests. As a result an unusually high percentage of students in these schools were classified as "borderzone" or "dull normal," and only a few were identified as "superior" in intelligence. Intelligence tests given in all elementary schools in January 1920 showed tremendous variation in test results as well. The median intelligence test scores for the schools ranged from forty-eight to 109. The scores varied by city neighborhood and, by implication, according to ethnicity and social class.

In spite of revelations such as these, or perhaps because of them, the intelligence testing program expanded rapidly in Oakland. By 1922 Dickson reported that the total number of tests given had reached 30,000, and that a "three-track plan" of classes for bright, average, and slow students "prevails throughout the entire city." The next year Dickson explained the Oakland system in Mental Tests and the Classroom Teacher, reporting that over 200 teachers had been trained to administer the tests.²¹

By the middle of the decade the adoption of intelligence testing for ability grouping was virtually complete. Intelligence tests, as well as the newly developed achievement tests, were administered to students in nearly all grade levels on a yearly basis. In 1925 some 1000 teachers were involved in revising the curriculum for the three-track system of ability grouping. In a report prepared late that year Hunter reflected on the changes that had swept over Oakland during his tenure as superintendent:

To adapt instruction to individual and group needs and capacities is one of the hardest problems which the teacher and school administrator have to solve....Nowadays the school system that has made no beginnings in the development of its curriculum along scientific lines and in the classification and promotion of its pupils in accordance with their scientifically pre-determined capacity is held to be in no sense attempting to meet the great problems before the schools.

Oakland, he was proud to say, had been a pioneer in that effort.²²

The Case of San Jose

Like Oakland to the north, San Jose was also a pioneer in adopting intelligence tests for use in the public schools. A city of nearly 40,000 in 1920, San Jose was heavily immigrant as well; over half the residents were first or second generation immigrants. The story of measurement in San Jose illustrates how one school system used tests to solve the "immigrant problem." The history of the early years of intelligence testing in San Jose also brings into sharp relief the issue of race, for university researchers and schoolmen alike interpreted ethnic differences in mental test performance as scientific proof of the "inferiority" of immigrants, especially those from Southern and Eastern Europe and from Latin countries. Armed with intelligence test scores, schoolmen classified students by ability, one result of which appears to have been a tracking system that separated immigrants and native-born Americans.

The initial impetus for intelligence testing in San Jose came from Stanford and a group of graduate students associated with Terman. The first contact came in 1916, through an experiment Terman conducted using the Stanford-Binet to select applicants for the city's police and fire departments. In 1919 a Stanford student named Kimball Young began a study of the "mental differences in certain immigrants" under the direction of Terman and Cubberley. With the cooperation of San Jose Superintendent of Schools, Alex Sheriffs, Young gathered information on all twelve-year olds in the schools to "discover if possible some of the causes of the difficulty in the education of children of South European ancestry in our public schools..." and whether their poor performance was due to "their alleged handicap" or to a "lack of mental endowments." He tested all students using the Army Alpha and Beta

and compiled other data including teachers' estimates of intelligence and accomplishment, grade location, and parents' occupational levels.²³

Among the "serious effects upon the public schools" of the recent "incursion of immigrants," Young noted in his study, discipline problems, difficulties teaching children from different cultures, and retardation. He discovered that 42 percent of the students were retarded, or behind their grade level. Examining four selected schools--"A" and "B" representing schools with predominantly Northern European students, "C" and "D" representing schools with primarily Latin students--Young concluded that Latin students were the farthest behind. Retardation was nearly twice as high in the Latin schools as it was in the Northern European schools, according to Young's statistics:

TABLE IV. RETARDATION IN CERTAIN SCHOOLS
IN SAN JOSE, 1919-1920

School	No. Pupils	No. Retard.	Percent Retard	Percent Latins in Schools
Total San Jose	4939	2098	42.4	40.0
School A	652	177	27.1	16.1
School B	629	187	29.7	17.5
School C	770	399	51.6	81.4
School D	744	464	62.1	72.4

The study also showed that while the students in the Northern European schools were fairly evenly distributed through the grades, those in the Latin schools were clustered in the first few grades. Nearly one-quarter of the students in the latter schools were in the first grade, while only six percent were in the eighth grade, as Young's data shows:

TABLE V. COMPARATIVE RETARDATION BY GRADE IN SCHOOLS A, B, C,
AND D; ALSO PERCENTAGE OF PUPILS IN THE VARIOUS GRADES

1. FOR SCHOOLS A AND B TOGETHER

Grades	No. Pupils	Percent Retarded	Percent in Grades
I	212	13.7	16.5
II	150	22.0	11.7
III	149	32.9	11.7
IV	137	36.5	10.7
V	161	36.6	12.6
VI	169	36.6	13.2
VII	140	24.3	10.8
VIII	163	29.4	12.7
Total	1281		

2. FOR SCHOOLS C AND D TOGETHER

Grades	No. Pupils	Percent Retarded	Percent in Grades
I	335	42.7	22.1
II	253	56.8	16.1
III	193	61.1	12.7
IV	208	64.4	13.7
V	145	71.0	9.6
VI	168	61.9	11.1
VII	126	61.9	8.3
VIII	96	46.8	6.3
Total	1514		

Thus a disproportionate number of immigrants apparently never completed elementary school.²⁴

For Young an analysis of the data collected provided the answer to the retardation problem. The Latin pupils "are decidedly inferior to the American pupils", he said, when measured by teachers' estimates, general intelligence (he computed the mean I.Q. for Latins to be 83), and economic status. Asserting that the language handicap "does not in fact exist in the case of the children of South European descent, to the great extent imagined," he concluded that "the true difficulty is one of mental capacity, or general intelligence, which makes the Latins unable to compete with the children of North European ancestry in the mastery of the traditional American public school curriculum."²⁵

"For the school man now," theoretical considerations are inconsequential, he said. "What he wants to know is how and what to do today." To solve the problems in the San Jose schools Young called for a "new educational policy." He recommended that "standardized tests should be applied throughout the elementary schools" as a preliminary step toward regrouping students on the basis of ability. Noting that "violent reclassification cannot be brought to pass over night," he observed that "the logical places to commence in San Jose...are in the schools so overwhelmingly Latin in population" because "it is there that the largest numbers of backward are found." He recommended that "pedagogical," or achievement, tests be used as well and suggested that the work could be facilitated by the establishment of a research bureau. After testing, the schools should be reorganized into three tracks, for "the superior, for the average, and for the backward," and the instruction for each group modified. For Young remaking the curriculum was "the crux of the matter." The new curriculum would provide "training for occupational efficiency, technique, and habits of work," habits of "co-operation" and "efficient living," and appreciation of the arts. For the able, the curriculum would be enriched; for the less able, it would provide training in skills and vocational education. A differentiated course of study, he said, would "fit the levels of capacity of the pupils" to "their future requirements in adult society."²⁶

San Jose was receptive to Young's suggestions, apparently, for in a footnote to his policy recommendations he observed that "since the above was written a new superintendent has taken over the control of the San Jose schools, and so far as this community is concerned we may look for rather radical changes in school policy. The new executive is already carrying out some of the recommendations of this chapter." That executive was Walter L. Bachrodt.²⁷

The new superintendent, appointed in 1921, moved quickly to attack the difficulties in the San Jose schools. In his analysis of the problem, Bachrodt agreed with Young. In his view, the great number of foreign born--from thirty-six countries--created an additional "burden" for the schools because the "school problem is not only one

of teaching fundamentals but also of assimilation (sic)." The immigrants from Italy--by far the largest single group--presented a particular problem of "colonization" for they were represented in very large numbers in three of the nine elementary schools. Examining the retardation question, he found results similar to Young's. In the Italian schools, Gardner, Grant, and Washington, the rate of retardation was the highest in the district (43.9, 46.6, and 58.0 percent respectively). The problem was compounded, Bachrodt said, by statistics which showed that relative to other cities in California, San Jose had an unusually high percentage of school-aged children in the population and an unusually low tax base. In other words, San Jose faced a financial crisis because of immigration and rapid growth.

Bachrodt's prescription for a solution to San Jose's ills followed closely Young's "new educational policy." In his first annual report in 1922, he observed that "the new field of Mental Testing has shed very valuable light upon this whole question" of retardation. He introduced his constituents to Terman's work with the Stanford-Binet and reviewed the army experiment. Acknowledging the recent criticisms of intelligence tests, he pointed out that as "we are primarily interested in why children do or do not succeed in school," it is sufficient to know that "from thousands of cases it has been definitely proven that mental tests do test the kind of mentality that it takes to go thru school...."²⁹

By the end of Bachrodt's first year as superintendent, intelligence testing had already made a preliminary beginning in San Jose. All students had been tested in two elementary schools as well as in the entire eighth grade in all nine elementary schools for an experiment in placing high school entrants in ability groups. The mean scores of the group testing provided a stark contrast of relative intelligence test performance in the city's different elementary schools. As in the retardation study, the three Italian schools ranked at the bottom with mean I.Q.'s of 71, 70, and 57 respectively. (The highest mean score was 94.) For Bachrodt the handwriting on the wall was clear. "It is well to face the facts," he said, "since our statistics

absolutely prove it, that the children in our schools that are predominably (sic) foreign do not compare with the children of other schools either in progress thru school or in marks earned on tests of native ability."³⁰

Though intelligence testing in 1921-22 was only conducted on an experimental basis, the structure existed to expand the program rapidly. Most of the work had been conducted by the research department of San Jose State Teachers College, which had strong ties to Stanford. The eighth grade testing had been conducted by James L. Stockton, who received his Ph.D. in 1920 under Terman. He was joined at San Jose by John K. Norton, formerly of the Oakland Research Bureau, who co-authored with Dickson an article on the use of tests in eighth grades for ability grouping and who helped to write Mental Tests and the Classroom Teacher. When Norton left the bureau in 1921, he was replaced by James C. DeVoss, another Terman Ph.D.

The preliminary experiments in intelligence testing and the evidence of massive retardation led Bachrodt to plan a reform for San Jose schools that closely followed Young's "new educational policy." On the basis of eighth grade testing, he concluded that the "biggest problem that a progressive school system has to face is the re-adjustment of pupils into classes of like ability." He proposed to extend the ability grouping experiment, establish a junior high school, introduce a more systematic program of educational and vocational guidance, and expand offerings in vocational education. Thus in 1922 San Jose initiated a major reform of public education based on testing, classification, and guidance.³¹

The Case of Palo Alto

In Palo Alto the problems schoolmen faced were of a different order. The population in 1920 was just a little over 7,000, immigrants were relatively few, the community was homogeneous and the residents prosperous. Sitting in Stanford's front yard, Palo Alto was in many respects an excellent example of a "progressive" school district, one which quickly instituted educational innovations developed at the neighboring university. It was in this context that Palo Alto began its own

kind of experiment with the use of intelligence tests just before the First World War.

In 1916 William V. Proctor joined Terman in the Department of Education and soon began a series of experiments using psychological tests for educational and vocational guidance. During and after the War, he introduced the tests in nearby Palo Alto High School, demonstrating how they could be used. While the primary use of intelligence tests in large urban areas was for classifying great numbers of students, Palo Alto took a different tack. In Palo Alto tests were used for sorting students, but the accent was on steering students through the educational system and on directing them towards appropriate colleges and occupations.

Consulting with Terman, Proctor began his experiment with the assistance of three graduate students--John Norton, later of Oakland and San Jose, Sam Kohs, and Mary Bess Henry. He secured the permission from Palo Alto High School principal, W.H. Nichols, to examine 107 entering students with Stanford-Binet tests. The purpose of the experiment, he wrote, was to "discover whether it is practicable to use intelligence tests for the educational guidance and grouping of high school pupils according to ability." Analyzing the records of twenty-seven students who had dropped out, Proctor concluded that there was strong relation between "elimination of mental capacity," and wondered whether the situation could not have been avoided through more appropriate guidance and placement.³²

Since the experiment brought Proctor into frequent contact with Nichols, the principal sought his advice on students having academic difficulties. As a result of this ad hoc advising arrangement, Proctor was asked to recommend appropriate placement for all students entering the high school at mid-year 1918. Pleased with the results, Nichols expanded the program to include all students entering the high school in September of that year. Word of the experiment spread fast, for the principals of four other Stanford area high schools, "upon hearing of what was being done," asked to be included in the continuing experiment.³³

In conjunction with this original project, Proctor conducted another experiment to determine whether psychological tests could predict high school success, and

---22---

how they could be used in educational and vocational guidance. In the fall of 1917 he gave Army Alpha tests to 1,349 students in eight Bay Area high schools (Palo Alto, Oakland Technical, Oakland Central, Fremont, San Mateo, Redwood City, Mountain View, and Santa Clara). Follow-up studies confirmed earlier findings: students with lower intelligence test scores were more likely to drop out of school. From his study Proctor predicted that students entering high school who tested below normal would drop out within two years and that few such students would ever graduate. For Proctor the implications were clear. With the I.Q. information at hand, the principal could "plan the curriculums of his pupils more intelligently....There will be fewer failures; more pupils will remain to take work that is adapted to their needs and capacities; and the high school will be less open to the charge of catering only to the intellectual aristocracy among its pupils." Proctor thus concluded that educational guidance with intelligence tests was feasible.³⁴

He then explored the possible uses of intelligence tests in the broader domain of vocational guidance. Discussing an analysis of the Army tests in 1920, he noted "discoverable differences in the intelligence levels of workers in various occupations." Comparing the vocational aspirations of some 900 high school students with the occupational intelligence levels revealed by the Army Alpha, as well as with vocational opportunities derived from the Census, he concluded that there was an urgent need for guidance. Not only did the students aspire to occupations with average intelligence levels above their own, they sought professional and managerial positions in numbers far out of proportion to the distribution of workers presently employed. His advice: steer students away from the professions and towards the agricultural, industrial, and mechanical pursuits. Proctor concluded that "The mental tests, if conservatively employed, will increase the probability that the counselor will give really helpful advice," and urged that "better and more systematic educational and vocational guidance of high-school pupils" was essential to the well-run high school.³⁵

Like Dickson in Oakland, Proctor decided to export the lessons he had learned in Palo Alto to a national audience. In Educational and Vocational Guidance (1925) he argued that the increasingly complex nature of society and the occupational structure called for a program of testing and guidance as a means of "conserving human values." The school would steer the child "along those paths of self-realization and social service to which he is best adapted by reason of his mental, moral, social, and physical endowment." And in that function, intelligence tests would play a major role.³⁶

The experience of Palo Alto, San Jose and Oakland in the first quarter of the twentieth century thus helps to explain why intelligence tests were adopted so rapidly in the public schools and what the relationship was between the introduction of the tests and the development of tracking. The process of adoption was complex and often varied according to the nature of the school system. Three common threads, however, are woven into the history of testing in these cities. On the one hand, a network of newly-emergent professionals provided the technology, the leadership, and the manpower for the movement which advocated using intelligence tests for ability grouping. On the other, the problems schoolmen faced--mushrooming size, increasing diversity, and a changing function of education--created a need for an instrument that could help classify students in an objective fashion. Finally, the values of the time--especially those which stressed the importance of efficiency, the authority of science, the power of the professional, and the reality of nativism--helped many to see intelligence testing and tracking as a "progressive" reform. While the school systems studied had already begun to differentiate their curricula and overall structure before the development of intelligence tests, the tests themselves acted as a powerful catalyst for the rise of ability grouping. The invention of intelligence tests and the rise of ability grouping in California schools early in the twentieth century thus heralded a new role in society for schools as sorters.

Endnotes

1. Gertrude A. Hildreth, A Bibliography of Mental Tests and Rating Scales (New York, 1933); Frank Freeman, Mental Tests: Their History, Principles and Applications (Boston, 1939), p.3; "Cities Reporting the Use of Homogeneous Grouping and of the Winnetka Technique and the Dalton Plan," Department of the Interior, Bureau of Education, City School Leaflet No. 22 (Washington, D.C., December, 1926).

2. Cf., Committee of Ten, Report (New York, 1894); and Clarence D. Kingsley, "Cardinal Principles of Secondary Education," Department of the Interior, Bureau of Education, Bulletin, XXXV (Washington, D.C., 1918).

3. Lewis M. Terman, "Trails to Psychology," in Carl A. Murchison, ed., A History of Psychology in Autobiography, II, (Worcester, Mass., 1932), p.321; Terman, "Genius and Stupidity," Pedagogical Seminary, XIII (1906), 310.

4. Lewis M. Terman, "The Binet-Simon Scale for Measuring Intelligence," The Psychological Clinic, V (1911), 204-205.

5. Lewis M. Terman, "Survey of Mentally Defective Children in the Schools of San Luis Obispo," Psychological Clinic, VI (1912), 136-37.

6. Lewis M. Terman, "A Report of the Buffalo Conference on the Binet-Simon Tests of Intelligence," Pedagogical Seminary, XX (1913), 549-54; Terman, "The Significance of Intelligence Tests for Mental Hygiene," Journal of Psycho-Asthenics, XVIII (1914), 122-24; Terman, The Measurement of Intelligence (Boston, 1916).

7. Terman, "Survey of Mentally Defective Children," 132.

8. Lewis M. Terman and Robert M. Yerkes to the General Education Board, Jan. 23, 1919, Lewis M. Terman Papers, Stanford University Library.

9. Lewis M. Terman to Caspar W. Hodgson, The World Book Company, Oct. 15, 1923, Lewis M. Terman Papers, Stanford University Library.

10. Lewis M. Terman, "The Use of Intelligence Tests in the Grading of School Children," Journal of Educational Research, I (1920), 20-31; Terman, Intelligence Tests and School Reorganization (Yonkers, New York, 1922), 17-19.

11. Fred M. Hunter, "A School System for All of the Children of All the People," Report of the Superintendent of Schools, Oakland, California, 1917-18 (Oakland, 1919).

12. Annual Report of the Public Schools of the City of Oakland, 1903-04 (Oakland, 1904), 56; Annual Report, 1911-12, 6; Vinnie Hicks, "Report of the Work of the Child Study Laboratory in Reference to Subnormals," Annual Report, 1911-12, 70.

13. Ellwood P. Cubberley, Report of a Survey of the Organization, Scope, and Finances of the Public School System of Oakland, California (Oakland, 1915), 3.

14. Hunter, Report, 1917-18, 33.
15. Virgil E. Dickson, "The Relation of Mental Testing to School Administration, with Special Reference to Children Entering School" (Unpublished Masters Thesis, Stanford University, 1917).
16. E. Morris Cox, "Survey of Nationalities," in Hunter, Report, 1917-18, 34-38.
17. Virgil E. Dickson, in Hunter, Report, 1917-18, pp.225, 226, 217; Dickson, "The Relation of Mental Testing to School Administration" (Unpublished Doctoral Dissertation, Stanford University, 1919), p.99.
18. Dickson, "Relation of Mental Testing," (1919), pp.74, 76, 79, 84, 85, and 86.
19. Virgil E. Dickson, "What First-Grade Children Can Do In School as Related to What is Shown By Mental Tests," Journal of Educational Research, II (1920), 480; Dickson, "The Use of Mental Tests in the Guidance of Eighth-Grade and High School Pupils," Journal of Educational Research, II (1920), 609.
20. Virgil E. Dickson and J.K. Norton, "The Otis Group Intelligence Scale Applied to the Elementary School Graduating Classes of Oakland, California," Journal of Educational Research, III (1921), 108.
21. Virgil E. Dickson, "Classification of School Children According to Mental Ability," in Terman, Intelligence Tests, 34; Dickson, Mental Tests and the Classroom Teacher (Yonkers, New York, 1923).
22. Fred M. Hunter, Superintendent's Bulletin, Special Edition (Oakland, January 1926), p.12.
23. Lewis M. Terman, The Intelligence of School Children (Boston, 1919); Terman, "A Trial of Mental and Pedagogical Tests in a Civil Service Examination of Policemen and Firemen," Journal of Applied Psychology, I (1917), 17-29; Kimball Young, Mental Differences in Certain Immigrant Groups (University of Oregon, 1922), pp.3 and 94.
24. Ibid., p.18.
25. Ibid., pp.63-64.
26. Ibid., pp. 66, 67, and 97.
27. Ibid., p. 66.
28. Walter L. Bachrodt, Annual Report of the Superintendent of Schools, 1921-22, 18.
29. Ibid., 21.
30. Ibid., 31, and 22.
31. Ibid., 32, 112, 113, 120-21, and 126.
32. William M. Proctor, "The Use of Intelligence Tests in the Educational Guidance of High-School Pupils," School and Society, VIII (1918), 474 and 506.

33. Ibid., 507.

34. William M. Proctor, "Psychological Tests and the Probable School Success of High School Pupils," Journal of Educational Research, I (1920), 270.

35. William M. Proctor, "The Uses of Psychological Tests in the Vocational Guidance of High-School Pupils," Journal of Educational Research, II (1920), 535 and 546.

36. William M. Proctor, Educational and Vocational Guidance (Boston, 1925), pp. 8 and 17.